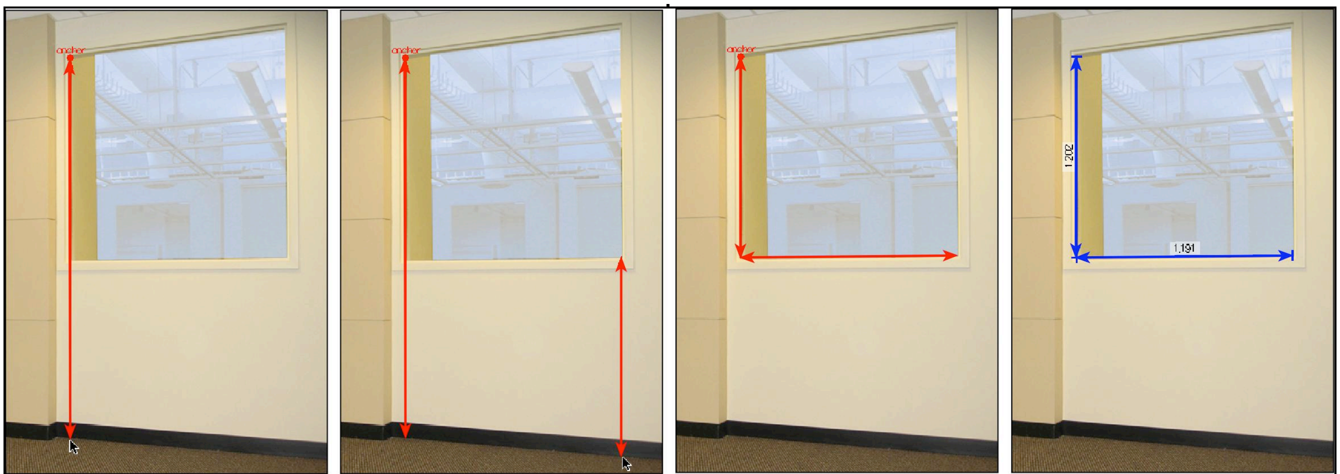


(a) (b) (c) (d)

Figure 1: The sequence shows how to measure a window with two mouse clicks. (a) The user clicks the top left corner of the window, drags down, and releases when the line touches the floor. During the drag, the line is constrained to remain vertical. The 3D location of the top corner of the window is computed and becomes the anchor (datum). (b) The user clicks the opposite, lower-right corner of the window and drags down to the floor. The 3D location of the opposite corner is computed. (c) PhotoMeter displays the perspective projection of the vertical and horizontal displacements between the anchor and the new point. (d) The user presses ENTER to keep this series of measurements and to annotate the image with the associated dimensions.

PhotoMeter: Easy-to-use MonoGraphoMetrics

Hendrik Müller and **Jarek Rossignac**

College of Computing, GVU Center, IRIS Cluster
Georgia Institute of Technology, Atlanta, GA, USA
mueller-hendrik@gmx.de, jarek@cc.gatech.edu

Abstract

*MonoGraphoMetrics is the science of computing 3D measures from a single image. Several recent research activities have produced theoretical principles and practical tools for extracting 3D measures from uncalibrated photographs. These tools require that the user identifies configurations of edges, which are used to establish constraints or to identify planes in the scene. The process involved is often laborious and its application is limited to images where the required configurations of edges are visible. In contrast, the work presented here is limited to photographs taken with a calibrated camera, oriented horizontally, at a known height above the floor. Under these conditions, a single mouse click provides enough information to compute the 3D position of any point **p** in the coordinate system of the camera, provided that **p** and its projection **f** on the floor can be identified in the image. With this approach, a novice user of our PhotoMeter system can easily measure dimensions and positions of windows, doors, pieces of furniture, and even people with one or two mouse clicks per measurement. The paper describes the geometric computation of the measurements, the user interface, and a study of how errors in the height and horizontal orientation of the camera affect the measurements.*

1 Introduction

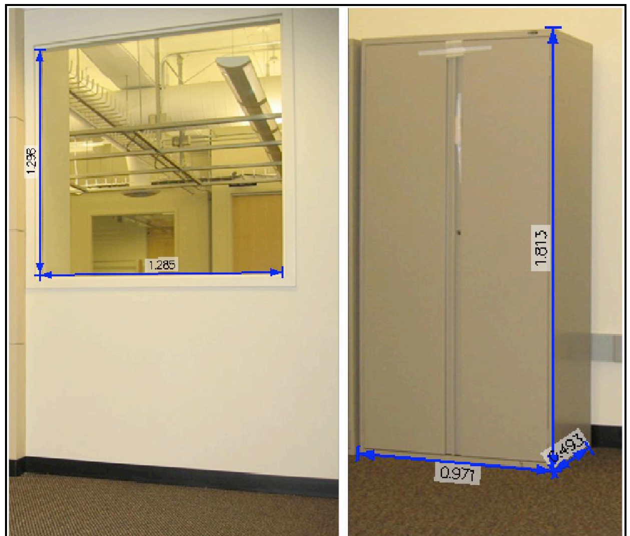
One of the primary challenges of Computer Vision and Computer Graphics is the automation of the creation of precise 3D models of real environments. The objective of the project described here is much more modest. We strive to provide human operators with a very simple-to-use and effective tool for performing real 3D measurements from a single photograph. Such a tool may be used for measuring indoor and outdoor spaces in order to obtain dimensions and positions of buildings, rooms, offices, windows, doors, furniture pieces, and even people.

The advantage of the proposed approach lies in its simplicity. A simple mouse operation, position-press-drag-release is sufficient to precisely measure the 3D location of a point visible in the image. Relative vertical and horizontal dimensions between such a user-defined 3D point and a previously defined point, serving as a temporary anchor, are clearly shown on the screen as mark-up arrows with dimension labels. This tool has been integrated within a complete interactive system, which we call PhotoMeter. The user of PhotoMeter can load a picture, perform the desired measurements, save the marked image for future references or email it to a colleague, client, or sub-contractor.

We envision applications of this technology in a variety of fields, involving architecture, real estate, interior decoration, and online purchase of furniture.

The simplicity and effectiveness of PhotoMeter result from a design decision, which restricts its use to pictures taken with a calibrated camera. More specifically, the user must know the horizontal field of view of the camera and the height of the camera above the floor or ground when the picture was taken. Furthermore, PhotoMeter relies heavily on the assumption that the camera was perfectly level (horizontal). Consequently, errors in camera calibration, in its height, and in its horizontal alignment will introduce errors in the measurements displayed by PhotoMeter. We provide an error analysis and suggest an approach for camera calibration, but recommend that PhotoMeter be used with tripods that permit to lock the camera tilt, hence ensuring a horizontal position and that make it easy to always set the camera at the same height. Alternatively, a simple stool may be used to support the camera.

The paper is organized as follows: In Section 2, we overview prior work on MonoGraphoMetrics, i.e., single view metrology. Techniques for camera calibration, camera leveling, camera height and lens distortion correction are explained in Section 3. The method used for computing the location of 3D points from a single mouse click is discussed in Section 4. The results from simulations and real data are provided and discussed in Section 5.



(a) (b)

Figure 2: PhotoMeter can measure the height and width of a window (a) or the height, width, and depth of a locker (b).

2 Related Work

There are particular areas of research that are applicable to our work: (1) camera calibration and lens distortion and (2) multiple- and single-view modeling, rendering, and metrology.

Devernay and Faugeras came up with a way of automatic calibration and removal of distortion from scenes of structured environments [DeFa01]. In reality, we cannot assume a pinhole camera model, where the projection of every line in space onto the camera is a line. We need to correct the image’s lens distortion. Devernay and Faugeras developed an algorithm that extracts edges from the image and measures how much these are bended compared to a straight line. On the basis of the degree of distortion, it is possible to correct the whole image. Finally, all lines become straight and the image can be considered as taken with a pinhole camera. This way, geometric computations can be easily performed with the image.

There have been various papers about image-based modeling/rendering and image-based metrology. These results are extremely related: metrology only adds the knowledge of a reference distance in the image of which basis other measurements can be obtained. For modeling purposes, multiple input images are necessary to get a 360-degree viewable model [DTM96]. For metrology, a single image can be enough [HAA97, LCZ99, CRZ00, Crim01, Crim02, KBB02, WW02, ElHa01, KCS03].

In 1996, Debevec et al. proposed a hybrid geometry- and image-based approach to model and render architecture from photographs [DTM96]. He developed a photogrammetric modeling method, which facilitates the recovery of the basic geometry of the photographed scene. Also, he came up with a better model-based stereo algorithm to recover how the real scene deviates from the basic model. Finally, with view-dependent texture mapping his method composites multiple views of the scene.

"Tour Into the Picture" presented in 1997 was one of the first approaches using a single image to create a 3D scene and an animation out of it [HAA97]. The user specifies a spidery mesh of the scene in the image by establishing the vanishing point. That way, perspective projection can be fitted and at most five planes (floor, ceiling, left wall, right wall, background) are established. After that, the user models the background, a foreground mask specifying all foreground objects, and also positions the camera. After this intensive user input, it is possible to move the virtual camera by changing its parameters and finally create a short walk-trough animation. These results do not create a real 3D model, but simulate a quite realistic walk-trough animation. In 1999, Criminisi et al. started to contribute tremendously to multiple- and single-view metrology and modeling [LZC99]. In contrary to Debevec's accomplishments, Criminisi's approach does not require multiple images, scene measurements, and the camera's internal calibration. Their algorithms use vanishing points and vanishing lines to establish planes trough defining a square with four control points. These are rectified and can now be used to measure on that plane knowing a reference distance of an object on the same plane. Furthermore, it was possible to compute internal camera parameters knowing three orthogonal vanishing points. This technique is restricted to exceptional images. Later, he added the possibility to measure between parallel planes and emphasizes the usage of uncalibrated images and finally the reconstruction of complete 3D scenes from single images [CRZ00, Crim01, Crim02]. He assumes to have images from which vanishing lines can be extracted. In [WW02, ElHa01], the authors describe extensions of Criminisi's method. First, the vanishing point based method is of equal precision and robustness compared with Criminisi's homography-based approach, but with less complexity [WW02]. Second, another approach neither needs vanishing lines nor calibrated images to create complete 3D models [ElHa01]. Finally, Kushal et al. suggested a method using two planes in the scene, selected from the user, to construct a three-dimensional representation of the image [KBB02]. Furthermore, he showed a method based on building the final model through high-level primitives like planes, spheres, cuboids, and so others [KCS03].

Although the solution proposed here is less general than the approaches mentioned above, it offers two significant advantages: (1) the user needs not waste time establishing a reference plane, since the height and orientation of the floor is known, and (2) measurements may be taken even when only a small portion of the floor below the measured points is visible, but no set of floor edges is available to define vanishing points (see Fig. 2 for example).

3 Preliminaries

When metrology is used for engineering or architectural applications, it is important to provide accurate measurements and to quantify the associated error. As in other MonoGraphoMetrics applications, the error in a measurement computed with PhotoMeter is due to the combined effect of several errors:

1) Camera calibration
2) Lens distortion
3) Horizontal alignment
4) Tripod height
5) Error in selecting a point in the image
6) Error in selecting its floor shadow

The cumulative effects of these errors on the accuracy of the measurements are discussed in Section 5. Here, we briefly suggest how to reduce these errors through proper camera calibration and image correction.

PhotoMeter uses the field of view of the camera. Test results have shown that specifications provided by camera manufacturers may be confusing or inaccurate. Therefore it is preferable to measure the actual field of view manually. Fig. 3 shows the starting panel in PhotoMeter explaining graphically how to measure the horizontal field of view by placing the camera parallel to a wall and reporting the perpendicular distance (Label "a" in Fig. 3) from the camera to the wall and the horizontal length (Label "b" in Fig. 3) of the visible portion of the wall. The horizontal field of view angle ("α") is computed automatically. To ensure that the camera viewing direction is perpendicular to the wall, we recommend placing the camera so that the left and right vertical edges of a wall, door or window appear perfectly flushed against the left and right borders of the image in the viewfinder. This camera calibration needs to be performed only once per camera.

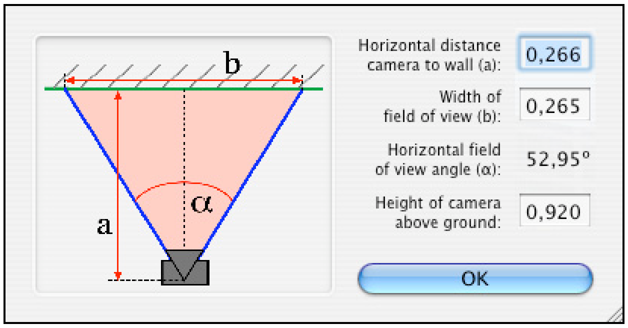


Figure 3: Initial camera calibration panel in PhotoMeter.

To ensure that the camera is perfectly horizontal when used to take a picture for PhotoMeter, one could simply place the camera on a horizontal surface, such as a stool or a tripod with fixed orientation and height. To ensure that the stool or tripod is perfectly horizontal, one could use expensive theodolites or a much cheaper laser level (Fig. 4), which includes a laser pointer that can be used to ensure that the height of the laser point on the wall is identical in several directions.

Figure 4: An inexpensive laser level for adjusting the horizontal.

The height of the camera must also be measured accurately. Pointing the laser at a vertical edge in the room may ensure that the height is measured vertically.

After the picture has been taken, distortions in the image caused by the lens of the camera should be corrected [DeFa01]. Several programs that correct lens distortion automatically or semi-automatically are available[1].

4 User Interactions & Computations

In this section, we explain how the user specifies the measurements on a single image and how the measurements are computed by PhotoMeter.

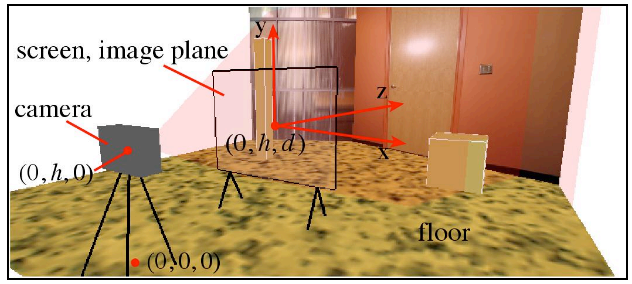


Figure 5: The coordinate system is defined by the camera.

Figure 6: The user interface of PhotoMeter. This screenshot also shows measurements that are already saved (blue arrows with dimensions) and a measurement series that is still in progress (red arrows with anchor).

To obtain the height h of a real 3D point **p** in the scene, the user clicks at the pixel **p’** where **p** appears on the screen, drags the cursor down to the floor, and releases the mouse-button at the pixel **f’** where the vertical projection of **p** onto the floor appears on the screen. To help the user ensure that the line from **p’** and **f’** is vertical, once **p’** is selected, the horizontal motions of the cursor are temporarily disabled, until the mouse-button is released.

As shown in Figs. 1 and 2, when **p’** identifies a point on a wall or on a vertical side of a furniture, the location of **f’** is obvious. Furthermore, when **p** lies on the floor, no dragging is necessary.

The first 3D point **p** identified during a series of measurements is used as an anchor (datum) for all subsequent measurements in the series. Subsequent 3D points may be added to the series until the ENTER key is pressed. PhotoMeter computes the vertical and horizontal displacements, in 3D, between each one of these points and the anchor. It overlays these displacements on the image by drawing their perspective projections as red lines. When ENTER is pressed, the series is frozen and the labels with dimensions added to the image. Then next 3D point will automatically become the anchor for the next series.

[1] An example for software that corrects lens distortion is LensDoc from Andromeda Software (http://www.andromeda.com/info/lensdoc/).

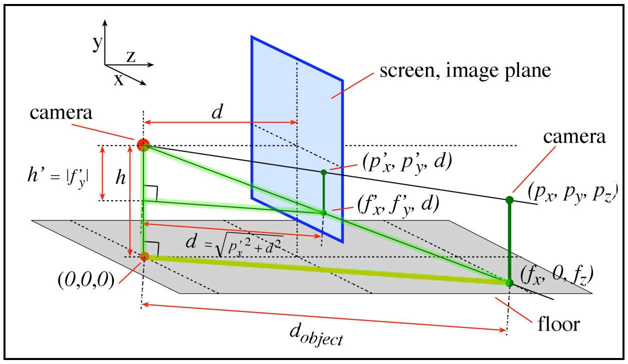


Figure 7: 3D view showing the geometric computation of the horizontal distance d_{object} from the object to the camera.

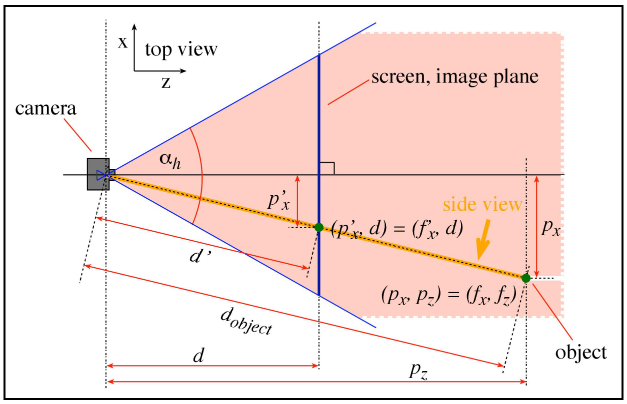


Figure 8: Top view showing how to compute the x and z coordinates of the object. The orange line defines a vertical section plane through the camera and the object.

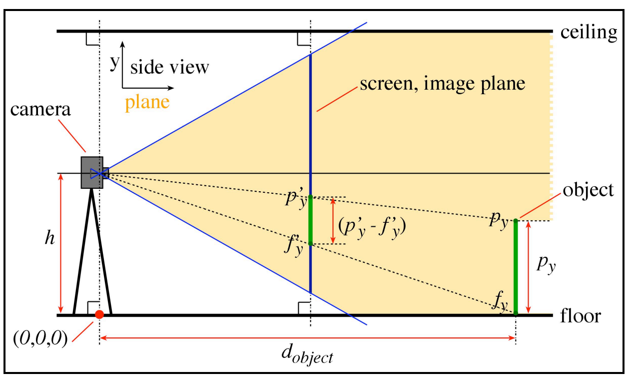


Figure 9: Side view of vertical section plane defined in Fig. 8 explaining the computation of the height of the object.

To find the coordinates $p = (p_x, p_y, p_z)$ of the selected 3D point **p** in the coordinate system of the floor projection of the camera, PhotoMeter uses **p’** and **f’**. (The origin is on the floor, below the camera and the camera’s focal center is at position $(0, h, 0)$. The formulae are provided in Fig. 10 and the computation illustrated in Figs. 7, 8, and 9. It takes as input the height h of the camera above the floor, the distance d from the camera to the image plane, the position $p' = (p'_x, p'_y)$ where the user clicked at in the image and the position $f' = (f'_x, f'_y)$ where the user released the mouse on the floor. The dragging constraint ensures that $f'_x = p'_x$.

$$p_x = \frac{p'_x \times h}{|f'_y|} \tag{1}$$

$$p_y = h + \frac{p'_y \times h}{|f'_y|} \tag{2}$$

$$p_z = \frac{d \times h}{|f'_y|} \tag{3}$$

Figure 10: Formulae for computing the 3D coordinates of a point **p** from the projection on the screen of the vertical line between **p** and the floor.

5 Results

The measurement errors due to camera tilt and height error are also a function of depth (i.e., distance between the measured object and the camera). These dependencies are illustrated in Figs. 11, 12, and 13.

Fig. 11 shows the acceptable horizontal camera tilt as a function of depth, so that the error in the object’s height does not exceed 10cm. A tilt of 1 degree would create an error of less than 10cm for an object located at 5m from the camera.

Fig. 12 shows the position error as a function of distance for a 1-degree tilt. For example, at a distance of 20 meters from the camera, a 1-degree camera tilt causes a position error of 35cm.

Fig. 13 plots the error in the position of **p** as a function of the error in camera height. The error depends on the height of the measured point. We have plotted the range of errors (dotted lines) and the mean error (red line).

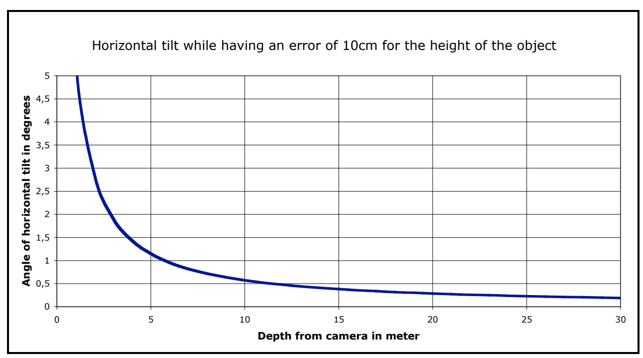


Figure 11: Acceptable tilt as a function of depth to ensure an accuracy of 10cm of the object’s height.

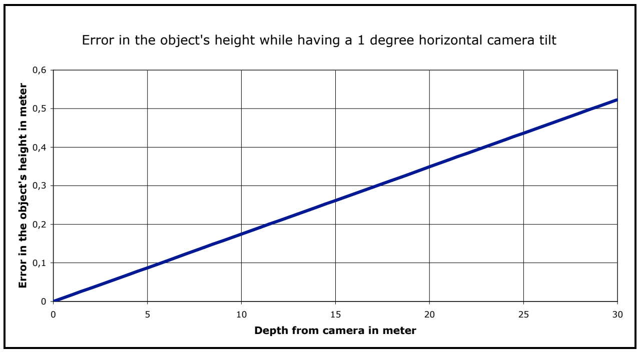


Figure 12: Error in computed height as a function of distance for a tilt of one degree.

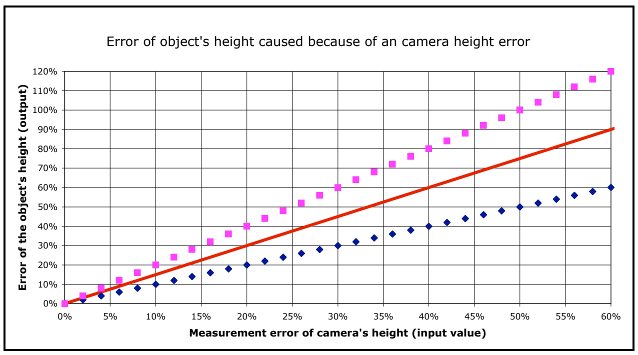


Figure 13: Position error as a function of inaccuracy in the estimation of the camera height.

A further source of error comes from the inaccurate selection of the points on the screen. It stems from the discretization of the image and from the difficulty of aligning the cursor with the correct pixel. Assuming that the user has selected the correct pixel, the maximal error of the position in the image would be the half of the diagonal of a pixel. Fig. 14 plots the maximal position error as a function of depth for points near the center of the screen and points near the edge.

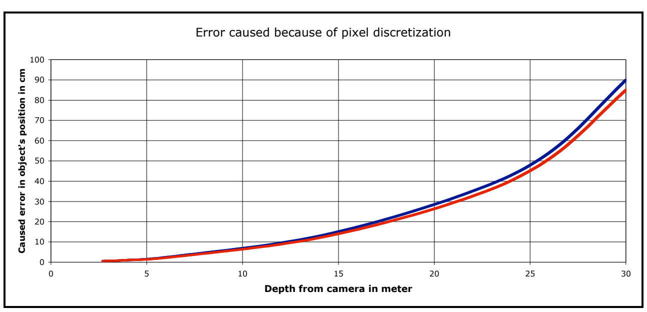
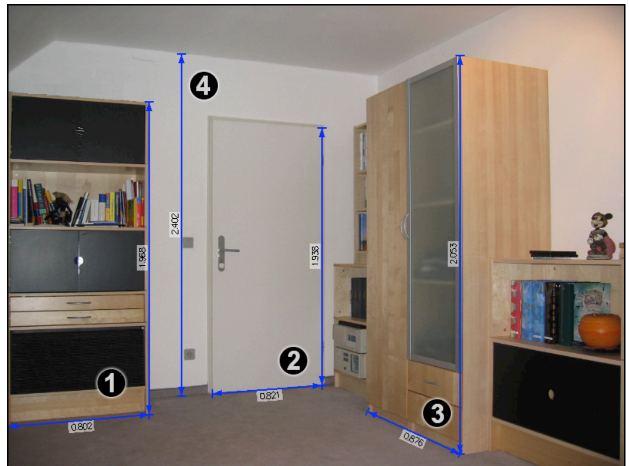


Figure 14: Error in computed height as a function of distance for due to pixel resolution for points near the center of the image (red line) and for points near the border (blue line).

To measure the cumulative effect of these errors, we have experimented with a variety of indoor scenes. The camera and tripod were calibrated as discussed in Section 3. In practice, we obtain an accuracy of less than 5cm when measuring the dimensions of windows, doors, or furniture pieces within 5m from the camera.



	Measurements 430-510cm away from the camera					
			Original	Photo-Meter	Error	
Objects	❶	height	199cm	197cm	1.0%	2cm
		width	80cm	80cm	0.0%	0cm
	❷	height	198cm	194cm	2.0%	4cm
		width	86cm	82cm	4.7%	4cm
	❸	height	201cm	201cm	0.0%	0cm
		width	93cm	88cm	5.4%	5cm
	❹	height	236cm	240cm	2.9%	4cm
				∅	3.1%	2.7cm

Figure 15: The error in the dimensions measured with PhotoMeter and shown here does not exceed 5cm.

6 Conclusion and Future Work

We describe a very simple-to-use tool for measuring vertical and horizontal distances between arbitrary 3D points from a single image, provided that the points and their vertical projection on the floor can be identified in the image.

The errors resulting from the cumulative effects of camera calibration, lens distortion, camera alignment, and pixel selection increased with distance from the camera. At 5m, we usually achieve an accuracy of less than 5cm.

We believe PhotoMeter as a viable alternative to physical measurements or to more elaborate photometry techniques for a range of applications where speed and ease-of-use are important and where the inherent inaccuracy is acceptable. Such applications include area floor and wall estimations for interior decoration and planning the layout of office or kitchen furniture.

In the future, we plan to explore how higher-resolution and local image-processing techniques can improve the accuracy of the manual point selection on the screen.

7 Acknowledgments

We wish to thank Alexander Powell for his guidance and help in developing the user interface for PhotoMeter.

We also wish to thank Lawrence Ibarria and Francisco Palop for their suggestions and support.

8 References

[Crim01] A. CRIMINISI. Accurate Visual Metrology from Single and Multiple Uncalibrated Images. *Distinguished Dissertation Series*. Springer-Verlag London Ltd., Sep 2001. ISBN: 1852334681.

[Crim02] A. CRIMINISI. Single-View Metrology: Algorithms and Applications. In *Proc. DAGM 2002 Symposium*, Zürich, Switzerland, Sep 2002.

[CRZ00] A. CRIMINISI, I. REID, AND A. ZISSERMAN. Single view metrology. IJCV, 40(2): 123-148, 2000.

[DeFa01] F. DEVERNAY AND O. FAUGERAS. Straight lines have to be straight: Automatic calibration and removal of distortion from scenes of structured environments. Machine Vision and Applications, 1, pages 14-24. 2001.

[DTM96] P. E. DEBEVEC, C. J. TAYLOR, AND J. MALIK. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *Proceedings*, *ACM SIGGRAPH*, pages 11-20, 1996.

[HAA97] Y. HORRY, K. ANJYO, AND K. ARAI. Tour into the picture: Using a spidery mesh interface to make animation from a single image. In *SIGGRAPH*, pages 225-232, 1997.

[ElHa01] S. F. EL-HAKIM. A flexible approach to 3D reconstruction from single images. *SIGGRAPH '01 Sketches and Applications*, 2001.

[KCS03] A. M. KUSHAL, G. CHANDA, K. SRIVASTAVA, M. GUPTA, S. SANYAL, T. V. N. SRIRAM, P. KALRA, AND S. BANERJEE. Multilevel modelling and rendering of architectural scenes. In *Proc. EUROGRAPHICS*. 2003.

[KBB02] A. M. KUSHAL, V. BANSAL, AND S. BANERJEE. A simple method for interactive 3d reconstruction and camera calibration from a single view. In *Proceedings Indian Conference in Computer Vision, Graphics and Image Processing*, 2002.

[LCZ99] D. LEIBOWITZ, A. CRIMINISI, AND A. ZISSERMAN. Creating architectural models from images. In *Proceedings EUROGRAPHICS*, pages 39-50, 1999.

[WW02] G. H. WANG, Y. H. WU, AND Z. Y. HU. A Novel Approach for Single View Based Plane Metrology. In *Proc. International Conference on Pattern Recognition*, vol.2, 556-559, Quebec City, Canada, Aug 2002.